Center for Independent Experts

Consultant Report on:

Groundfish Assessment Review Meeting III: Modeling

February 25 – March 1, 2008 Woods Hole, MA

Paul A. Medley
Sunny View
Jack Hole
Alne, YO61 1RT
UK
paul.medley@virgin.net

Contents

Executive Summary

This report presents additional comments on the terms of reference to those in the panel consensus report. The consensus report indicates the modelling approaches which were considered the best available. While problems were identified with these assessments during the meeting, clear solutions are not available at this time.

The following comments in this report are potential avenues for improving the current assessment models, for example reducing problems indicated by retrospective patterns:

- Explore alternative models describing natural mortality (e.g. M inversely proportional to length, or based on ecosystem models or on multispecies VPA, allowing M to change over time).
- Separate sex specific population models where growth differences exist between males and females.
- Consider the local temperature (and other measurable variables) effect on survey catchability.
- Separate catches by fleet and gear type, estimating separate selectivity patterns, producing a more functional description of fishing rather than statistical description of the data.
- Develop a Bayesian approach to fitting biomass dynamics models rather than rely on relative trends in survey indices (AIM).

Risk analysis should become the standard approach for dealing with uncertainties that cannot be resolved through available data and/or future scientific research. In many cases, there will be insufficient data to resolve issues. It should be necessary to develop a standard method to present risk analysis so that decision-makers can understand this dimension of uncertainty.

Background

The Groundfish Assessment Review Meeting (GARM) is a peer review of stock assessments of 19 important groundfish stocks that are managed by the New England Fishery Management Council. GARM-III is taking place in 2007-2008, consisting of four meetings, where the successive meetings incorporate methods and results that were accepted at previous GARM-III meetings. This report covers the second meeting in the series.

The first "Data" Meeting (October 29 – November 2, 2007) reviewed the commercial and survey data that will be used in the stock assessments. This report describes the second review "Modeling" Meeting (February 25 – 29, 2008), which intends to determine the most appropriate stock assessment methods and models for each of the 19 stocks. After this there will be a "Biological Reference Point (BRP)" meeting (April 28 – May 2, 2008) to update or redefine BRPs for each of the 19 stocks and a final meeting (August 4 - 8, 2008) to use all of the methods proposed from the previous three meetings, along with updated data, to estimate historical and current fishing mortality rates, biomass and status for each stock.

Review Activities

Review documents were received a week before the meeting (Annex I). The meeting took place in Woods Hole 25th February – 1st March 1 2008. The panel was required to determine and write down its viewpoint on the quality and soundness of the science, methods and data with regard to each Term of Reference (see Annex II). Among other things, the panel was required to consider whether the data are adequate and used properly, the analyses and models were appropriate and the conclusions are reasonable.

The panelists were required to produce a consensus report and, in addition, the CIE panelists are required to produce individual reports representing their own views. There were three CIE reviewers at the meeting: Cynthia Jones, Jose De Oliveira and Paul Medley. This report is the individual CIE report of Paul Medley. It does not repeat consensus views on the terms of reference and the 19 stock assessments, but does state the independent views of this reviewer on issues arising from the meeting.

The following presents additional comments on the terms of reference to those in the panel consensus report. The consensus report indicates the modelling approaches that were considered the best available at this time. The comments in this report are potential avenues for improving the current assessments. While problems were identified with the assessments, clear solutions are not available at this time. Further research outlined here should reveal better approaches in future.

ToR 1

For each stock, consider the applicability of one or more of the following modeling approaches to assess stock status:

- **Index methods**
- **Production Models**
- **Age- or Length-based Models**

Summary of Findings

The relative index based method (AIM) needs an assumed population model to define the biological reference points (BRP). The relative index approach makes it appear fewer assumptions are made, but to some extent this is a result of assumptions being more obscure rather than not being present. The alternative surplus yield model approach suggested at the meeting would not deal with this issue unless a relationship between stock size and stock growth is proposed.

Biomass dynamics models are not currently used in the proposed assessments and in a number of cases have been rejected during previous reviews because they have been unable to explain the observations adequately. Survey data may be suspect as an index of abundance for the stocks that are not primary target stocks and have no age data. The survey in these cases may not be a good index of abundance (Fig. 1).

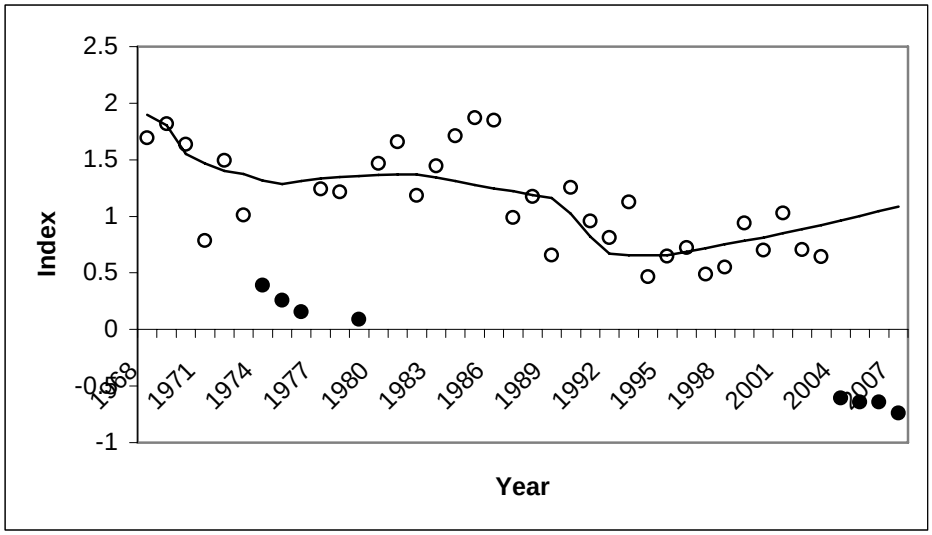

Figure 1 Ocean Pout survey index and fitted biomass dynamics model to selected indices (○). If all data were included in the series, no maximum likelihood model could fit the data. Selective removal of some years' data (●) allowed a fit. The removed data are not consistent with the series and call into question whether the index is valid in all years. If the index is invalid, switching to an index only model will not fix this problem. In particular, the last four index points for this species may contain no information for this stock.

Conclusions and Recommendations

Biomass dynamics models should be used in preference to relative index models. Their assumptions are explicit and combine biological reference points and status estimates. They can provide a reasonable empirical description of the behaviour of a population in response to fishing. They do not provide an accurate biological description of the population however. Therefore the advice that they provide can be

simplistic and they do not cope well with changes in productivity (e.g. reductions in recruitment due to climate change).

Because the survey data may be poor for some species, it may be justified to treat some data as missing or use a robust likelihood to down-weight poor index data points. If data points are removed, some justification is required based on how the survey was conducted.

Given the previous experience, ASPIC is not an adequate approach, and alternative Bayesian approaches should be developed. It would be useful to develop a method to define informative priors based on, for example, ecological factors, ecosystem models and so on. For example, individual point estimates from ecological simulations and publications of parameters or functions of parameters (e.g. MSY) can be used to generate prior probability densities for Bayesian assessments (Silverman 1986). MCMC does not always work well for biomass dynamics models, and alternatives (SIR or rejection algorithm; Gelman *et al.* 1995) may prove more robust.

ToR 2

For certain stocks that are aged, compare and contrast the utility of statistical catch-at-age vs. VPA based models with respect to the following criteria:

- **Retrospective patterns**
- **Flexibility to account for alternative parameterizations**
- **Ability to incorporate external sources of information, especially tagging and environmental data**
- **Ability to estimate parameters incorporating prior, external information.**

Summary of Findings

As was made clear throughout the meeting, virtual population analysis (VPA) and statistical catch-at-age (SCAA) are very similar approaches. Both apply the same population model, but have different assumptions with respect to errors in the catch data. SCAA models are more rigorous in their treatment of error, and it is likely that development of modelling approaches will advance in this area rather than VPA. However, treating the catch-at-age as known exactly does make fitting easier and can improve results if SCAA assumptions regarding catch errors are poor. Allowing selectivity to change from year to year and having a low error on catch and age composition data in SCAA should produce very similar results to VPA.

If developing SCAA models, it may be better to use software developed locally, such as ASAP. This should support adaptation to local needs. SS2 generally requires external support which may not always be available. SS2 uses standard numerical software which could be used by any other stock assessment software, so it does not offer any other significant advantages. Other software, such as CASAL, could equally well be used, but suffers the same problem of distant support.

Length-based selectivity may give better more parsimonious results in some cases. For example, different sex selectivity might be explained by different growth forms without reference to sex-specific selectivity.

The Baranov equation may fit better at higher fishing mortality because it takes better account of within-year depletion. This can be useful when accounting for date of a survey within the year and/or when fitting commercial CPUE. Whether the Baranov equation is worth the additional numerical overhead can be tested using goodness-of-fit statistics, as the data and number of parameters remains the same as for Pope's approximation.

Conclusions and Recommendations

In the longer term, SCAA models should be preferred to VPA. Because of the way SCAA models are set up and fitted, they can more naturally incorporate external data and prior probabilities on parameters and allow greater flexibility in terms of exploring assumptions. On balance therefore, SCAA approaches are preferred, but switching to SCAA when a VPA is producing good results is not a priority, but a long term aim.

SCAA software should be chosen for ease of use and local support. Software will need to fit local model, data and output requirements. Ideally the software should allow length as well as age-based selectivity, separate sex models, and for selectivity to change over time from blocked periods, penalised random-walk through to free selectivity (like VPA).

ToR 3

Address the implications of zeros in the evaluation of fishery independent indices.

Summary of Findings

The issue of dealing with zero observation when attempting to use log-normal likelihood was not found to have an important influence on stock assessments. The current default of treating values as missing is clearly better than having a fixed constant value.

Zero observations suggest that a likelihood which implies zeros are impossible is inappropriate. However, survey data suggest that the log-transform stabilises the variance of the series. Alternative Box-Cox transforms and alternative likelihoods have not been tried.

In some cases, zero observations can be removed by combining age groups. In addition, an SCAA approach where a zero observation may be explained using the multinomial, can eliminate the problem.

Conclusions and Recommendations

Clearly, treating zeros as "no observation" does not currently produce problems, but could under some circumstances lose important information on the depleted state of a stock. In the long term, the problem is likely to disappear as assessments change to an SCAA approach and a more appropriate likelihood. In the mean time the following alternatives could be explored:

- Use alternative Box-Cox transform. Square root may not produce a constant variance (equivalent to Poisson), but higher roots (e.g. 4 or 8) may do. There is no theoretical justification for this approach, however.
- Apply the quasi-Poisson likelihood. Alternative likelihoods that allow zeros, such as the negative binomial, are more difficult to fit as they are not scale invariant. The Poisson can be fitted using iterative least squares, but there is still increased numerical overhead.
- Use a veil-line approach to model the likelihood for zeros, as integral over an interval rather than a point value. Based on the sampling effort (number of tows), the probability of catching nothing can be estimated using the log-normal as the integral of x values from 0 to 1.0 (equivalent to an integral of the normal from 0 to negative infinity), which can be calculated easily using the cumulative normal and the current estimates of the mean and CV.

ToR 4

Examine potential factors responsible for retrospective patterns.

Summary of Findings

A significant retrospective implies that the biomass estimate is poor. However, trends may continue to be valid and can still be used for management advice. Where controls deal with absolute biomass (i.e. TACs) this may continue to be a problem. However, analyses presented at the meeting on this topic were not exhaustive. For example, there were no simulations testing for the effect of gradual changes in survey catchability, natural mortality or catch reporting, or other factors that represent mis-specification in the population models.

The most likely cause of the observed retrospective patterns is changes in the survey catchability. The alternatives, a change in natural mortality or unreported catch, seem less likely based on the information presented. A discrete change in natural mortality is unlikely without corroborative evidence, such as observations on diseased fish. There was no evidence at this meeting of changes in unreported catch or discarding. There was no evidence of an "observer effect", which could result in a change in behaviour at sea being missed. However, catch errors should not be ruled out, as there was also no information that would eliminate this possibility.

It is possible that natural mortality has changed in line with the length-at-age. To explore this, a simple VPA was fitted to the available catch-at-age and survey index data for Gulf of Maine cod presented at the meeting, applying a natural mortality that is proportional to the reciprocal of the length (Lorenzen 2005). This model of natural mortality has been proposed based on observations on small fish mortality and restocking. As a function, it may not be accurate, but may represent natural mortality better than assuming it is constant, as smaller fish are thought to be more vulnerable to predation.

The effective natural mortality during a unit time, taking into account von Bertalanffy growth (L_∞, K), start length (L_t) and natural mortality at unit length (M_1), is:

$$M = Ln\left(\frac{L_t}{L_t + L_\infty\left(e^K - 1\right)}\right)\left(\frac{M_1}{L_\infty K}\right)$$

Using this equation allowed the natural mortality in the VPA to be adjusted for the smoothed observed length-at-age, which has changed over time. Observed initial length (L_t) and K were determined by the smoothed mean lengths. M_1 was chosen to give a natural mortality of around 0.2 year^{-1} for the oldest age groups at the start of the series, and asymptotic length was fixed at 148cm (www.fishbase.org). In this exploratory analysis, these gradual changes in natural mortality appear to produce less of a retrospective pattern (Fig. 2), but patterns are not eliminated. This alternative population model did fit the observations much better (as indicated by the sum-of-square residuals between the observed and expected age specific survey cpue – SS = 256 vs. 1190).

It is not clear what effect the lack of contrast in the data may have on the retrospective problem. It was indicated at the meeting that initial exploration suggested that this was not a problem. Nevertheless, a stock assessment attempts to match observed catches against the depletion in relative indices, and thereby estimates the total biomass. A lack of good contrast in the data (variation in catch) may prevent good reliable estimates of biomass and prevent proper treatment of the problem.

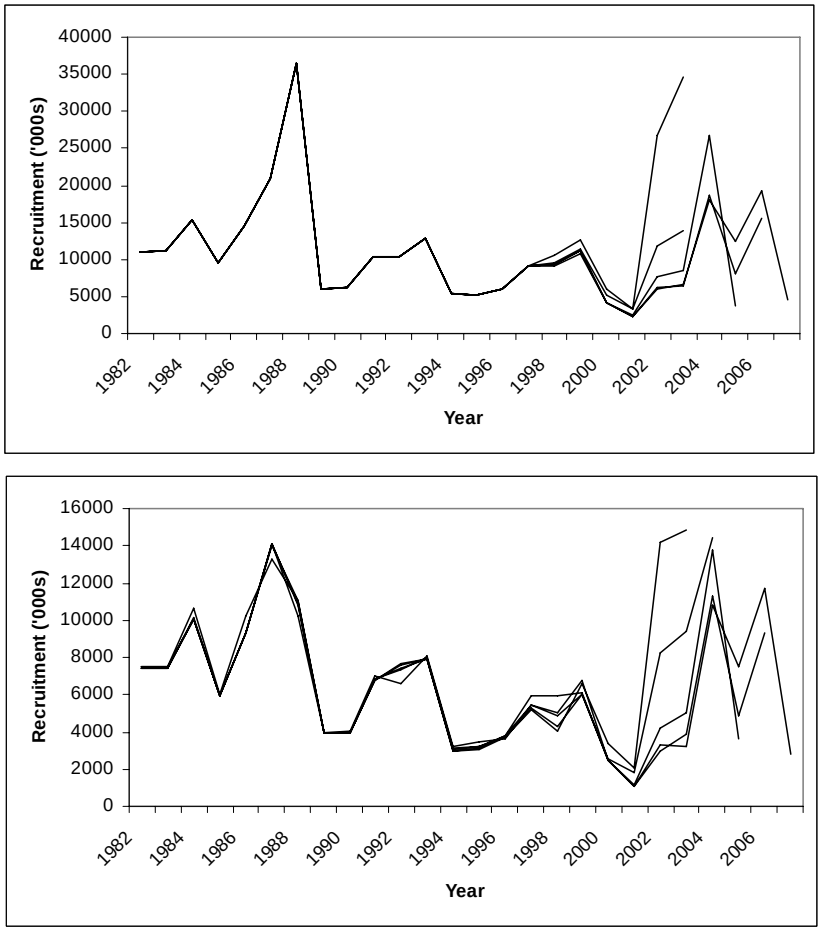

Figure 2 Retrospective pattern for Gulf of Maine cod VPA with natural mortality inversely proportional to length (top) and fixed natural mortality (bottom).

Conclusions and Recommendations

Preliminary results suggest the use of the natural mortality inversely proportional to length may be worth exploring as a way to improve the population model. This may explain the observations better and reduce retrospective patterns. More generally, this result suggests that retrospective patterns and the model might be improved by changing the model specifications to reflect more accurately the population dynamics.

It is likely that the main cause of severe retrospective patterns is changes in survey catchability, with the second possibility being changes in catch reporting and discarding. Without alternative information, splitting survey time series should be the default correction applied if a single model is required. Commercial CPUE should be used to check the validity of a split, even if not used to fit the model.

There appear to be significant problems with the reliance on surveys covering so many species. Further research on the survey methodology, gear efficiency and selectivity appears warranted. Research in Alaska indicates local temperature has an important influence catchability of flatfish (Spencer *et al.* 2007). Such physical oceanographic data should be available for these surveys, and provide a start point for investigation.

ToR 5

For each stock, define the assessment model that will be used to determine stock status and productivity characteristics until the next "benchmark" assessment is conducted. Where possible, apply the models to data (probably through 2006), to obtain current and historical estimates of F and B and estimates of uncertainty.

Summary of Findings

Recommendations on the individual stocks are made in the panel report and are not repeated here and general recommendations on the modeling approach are made in ToR 1. The following additional notes apply across a few stocks as indicated.

There was considerable discussion on domed selectivity (mainly Gulf of Maine cod), which seems to be a common issue in age structured stock assessment. This effect may be due to true selectivity changes, unrecorded mortality or model misspecification. There was insufficient evidence presented at the meeting to choose between these as the major cause in the stocks reviewed.

As well as being a true representation of selectivity, domed-selectivity may be another indication of model mis-specification. Preliminary modelling of natural mortality as a decreasing function of length produced less domed-selectivity than a constant natural mortality (see ToR 4), the opposite of what might be expected. Clearly, an alternative explanation for domed-selectivity would be flat-topped selectivity and natural mortality varying with age. Domed selectivity may turn out to be an artefact of errors in the way the population is modelled and the corrections produced by the model fit that tries to account for this structural error.

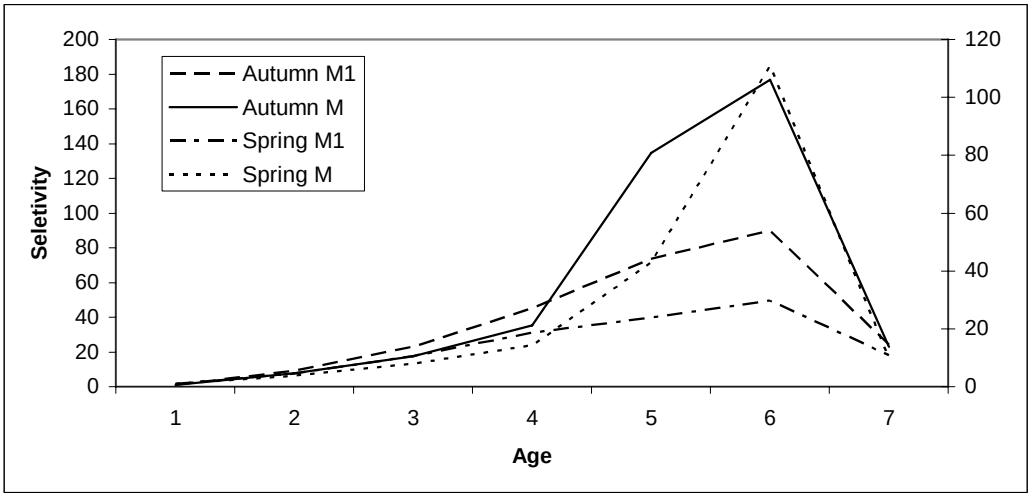

Figure 3 Selectivity pattern (mean F at age) for VPA catch-at-age models with natural mortality inversely proportional to length (M1) and fixed natural mortality (M) for the autumn and spring surveys of Gulf of Maine cod. The selectivity shows less "domedness" when natural mortality decreases with length.

The assessments without age data were clearly weaker than assessments based on catch-at-age. Ocean pout was one of the low-data, low value stocks presented. There were no age data, but there were length frequencies from the survey and observer program, and discards had been estimated where possible. There were no biological reference points produced from the AIM assessment applied. There appear to be problems with the survey index for this species and a few outlier points have prevented a maximum likelihood fit of a biomass dynamics model (ASPIC).

Conclusions and Recommendations

In some cases, the stock assessment will be improved through more accurate population dynamics models rather than using adjustments to standard models. These include the stocks where concern was expressed over retrospective patterns and domed-shaped selectivity. The following areas are suggested as potential areas of development which may improve various assessments:

- Separated sex specific population models for flatfish and other species where growth differences exist between males and females.
- Local temperature (and other measurable variables) effect on catchability, particularly for flatfish. Pilot analysis could simply look at correlations between mean bottom temperature and catch rates on each tow within the survey (ToR 4).
- Separate catches by fleet, estimating separate selectivity patterns. Gillnets and trawls are likely to have very different selectivity. The results would then represent a more functional description of fishing rather than statistical description of the data.

Where possible, observation error for the different survey indices should be estimated within the fitted model. In some cases this may be difficult or not possible, and a minimum-maximum error range or fixed error is required to weight an index series correctly. A non-parametric smoother can be used through indices to get weights, and

the residuals used to estimate the error. The degree of smoothing should depend in part on the maximum age (natural mortality). This is appropriate because any population model will act like a smoothing function through a biomass index, so this is the best (lowest) error which could be obtained from a fitted model.

The index based assessments (AIM) should be moved to assessments based on population models. Biomass dynamics based assessment would be the most robust approach, particularly if able to use additional ecological and other information, with the catch and CPUE time series, to improve parameter estimates, as could be developed under a Bayesian framework. A Beverton and Holt estimator of mortality that requires only mean length (Gedamke and Hoenig 2006) could also be used for comparison. However, given the change in length at age observed in a number of species, it may be considered too unreliable.

Risk analysis should become the standard approach for dealing with uncertainties that cannot be resolved through available data and/or future scientific research. In many cases, there will be insufficient data to resolve issues. It should be necessary to develop a standard method to present risk analysis so that decision-makers can understand this dimension of uncertainty.

As done under the STAR process, the scientists responsible for the assessments could be asked to consider the main axis of uncertainty, and present results illustrating the best fit and some scenarios across this axis. Communicating this uncertainty is most easily done through decision tables. However, for decision tables to be effective, not only the states of nature, but also the range of possible actions needs to be defined as well as measures of outcome. Management actions could be the range of possible target fishing mortalities defined in the harvest control rules for the different states of nature, and the outcomes could be the resulting future states of the stock.

ToR 6

Evaluate the sufficiency of the assessment models to estimate measures of stock status consistent with Biological Reference Points.

Summary of Findings

AIM cannot estimate biological reference points. These need to be provided from population models that may not fit the data, with the result that these assessments may be inconsistent. For example, the ocean pout assessment uses reference points from a rejected model, whereas the assessment is based on the index only. Given that the index is likely to be the reason the population model cannot fit the data, this does not really solve the problem.

A concern with using biomass dynamics models (ASPIC) was that biological reference points changed from year to year. This need not be the case, and the models can be refitted conditional upon relevant reference points (B_{MSY}, MSY) being fixed values. Reference points can then be updated when required in the usual way.

Conclusions and Recommendations

NEFSC should develop software for fitting biomass dynamics models that implements Bayesian fitting methods, allows the inclusion of priors and controls whether reference points are allowed to change or not. Fixing the reference points at a

particular value will need to be considered as a source of uncertainty, although in most cases it is unlikely to be the main source. It may be necessary to identify a method to reject individual annual index data points as uninformative (Fig. 1), and provide informative priors on important parameters or parameter functions (e.g. MSY), through ecosystem models, for example.

Comments on Review Process

Nineteen was a large number of assessments to review, and therefore none could be looked at in depth. As a result the review of each stock was fairly superficial, and there was considerable reliance on experts already familiar with the fisheries and stocks. More in depth review of some of the assessments from this multispecies assemblage to guide overall management may be valuable. However, given the constraint above, the terms of reference were met.

References

Gelman, A., Carlin, J.B., Stern, H.S. Rubin, D.B. (1995) Bayesian Data Analysis. Chapman and Hall, New York.

Gedamke, T. Hoenig, J.M. (2006) Estimating Mortality from Mean Length Data in Nonequilibrium Situations, with Application to the Assessment of Goosefish. Transactions of the American Fisheries Society 135:476–487, 2006

Lorenzen, K. (2005) Population dynamics and potential of fisheries stock enhancement: practical theory for assessment and policy analysis. Philosophical Transactions of the Royal Society of London. Fisheries Theme Issue 2004.

Silverman, B.W. (1986) Data estimation for statistics and data analysis. Monographs on Statistics and Applied Probability 26. Chapman and Hall, London.

Spencer, P., Wilderbuer, T., Stockhausen, B., Ianelli, J. (2007) Estimation of temperature-dependant catchability for eastern Bering Sea flatfish. Alaska Fisheries Science Center, National Marine Fisheries Service, 7600 Sand Point Way NE, Building 4, Seattle, WA 98115. Current Research: Personal Communication.

Annex I Bibliography

TOR #1	
WP 1.1	Data summary 19 stocks
WP 1.2	History of Model Selection 19 stocks
WP 1.1 & 1.2	Section M: Pollock - corrected 2-22-08
WP 1.3	NE Fishery Observer Program Maps
TOR #2	
WP 2.1	Statistical Catch at age models
WP 2.2-a	VPA vs ASPM - Gulf of Maine cod
WP 2.2-b	ASPM retrospective - Gulf of Maine cod
WP 2.3	ADAPT VPA vs ASAP models - Gulf of Main cod
WP 2.4	Simulated data sets
WP 2.5	ASPM - GB yellowtail flounder
TOR #3	
WP 3.1	Filling Zeros
WP 3.2	zeros_2006
WP 3.3	ICES Report - zeros
TOR #4	
WP 4.1	Report of the Retrospective Working Group
TOR #5	
WP 5.1	Model selection - Initial Recommendations
for WP 5.1	Table 5.1 of WP 5.1
TOR #6	
WP 6.1	Sufficiency of Models for BRPs
Supplementary Papers	
TOR 1	Gulf of Maine Winter Flounder - appendix 1 Scale run
TOR 2	Yellowtail: dome-shaped selectivity
Background	
GARM III	Data Meeting Report December 2007
GARM II	2005 Report
NEFSC 2002	CRD02-04 BRPs
NEFSC 2002	CRD02-12 Cod TRAC
TRAC 2005 01	GB YT Flounder

Annex II Statement of Work

Statement of Work for Dr. Paul Medley

External Independent Peer Review by the Center for Independent Experts

GARM-III "Models" Meeting:
Statement of Work (SOW) for CIE Panelists
(including description of GARM-III Chairman's duties)

General

The Groundfish Assessment Review Meeting (GARM) brings together stock assessment experts to peer review work on the status of 19 important fish stocks that are managed by the New England Fishery Management Council. GARM-III takes place in 2007-2008, and it will consist of four meetings that are cumulative in nature (i.e., successive meetings incorporate methods and results that were accepted at previous GARM-III meetings). Each meeting will have a chair as well as external panelists. A brief description and dates of the four GARM-III meetings are given below:

1. "Data" Meeting (October 29 – November 2, 2007)
Review the commercial and survey data that will be used in the stock assessments. Identify appropriate statistical methods for analyzing those data (including bycatch and discard issues, changes in growth rates and other life history traits, issues related to merging databases, etc.). Other sources of data to be considered are tagging programs for cod and yellowtail flounder, and Industry-Based Surveys. Candidate sources of data relevant to ecological and ecosystem considerations will also be described.

2. "Modeling" Meeting (February 25 – 29, 2008)
Determine the most appropriate stock assessment methods and models for each of the 19 stocks. Perform runs of those models to obtain results (historical and current estimates of F and B) based on commercial and survey data, probably through calendar year (CY) 2006. Evaluate retrospective patterns and their importance for status determination.

3. "Biological Reference Point (BRP)" Meeting (April 28 – May 2, 2008)
Update or redefine BRPs for each of the 19 stocks. Use data available through CY2006. Consider whether the BRPs are reasonable in light of results from the "Modeling" Meeting. Define the appropriate initial conditions for forecasting and rebuilding strategies, particularly with respect to trends in biological attributes, recruitment and survival rates. Comment on relevant ecosystem considerations as they relate to rebuilding strategies.

4. GARM-III "Final" Meeting (August 4 - 8, 2008)
   Use all of the methods proposed from the previous three meetings, along with survey and catch information through CY2007, to estimate fishing mortality rates and biomass for each stock. Based on procedures from the BRP Meeting, finalize the BRPs, appropriate initial conditions, and biological assumptions related to forecasts. Determine the status of each stock.

This SOW applies specifically to the GARM-III "Modeling" Meeting, which will take place at the Woods Hole Laboratory of the Northeast Fisheries Science Center (NEFSC) in Woods Hole, Massachusetts, from February 25 - 29, 2008. The meeting will have a chairman (non-CIE) as well as external panelists, three of whom will be from the Center of Independent Experts (CIE).

Overview of CIE Peer Review Process:

The Office of Science and Technology implements measures to strengthen the National Marine Fisheries Service's (NMFS) Science Quality Assurance Program (SQAP) to ensure the best available high quality science for fisheries management. For this reason, the NMFS Office of Science and Technology coordinates and manages a contract for obtaining external expertise through the Center for Independent Experts (CIE) to conduct independent peer reviews of stock assessments and various scientific research projects. The primary objective of the CIE peer review is to provide an impartial review, evaluation, and recommendations in accordance to the Statement of Work (SoW), including the Terms of Reference (ToR) herein, to ensure the best available science is utilized for the National Marine Fisheries Service management decisions.

The NMFS Office of Science and Technology serves as the liaison with the NMFS Project Contact to establish the SoW which includes the expertise requirements, ToR, statement of tasks for the CIE reviewers, and description of deliverable milestones with dates. The CIE, comprised of a Coordination Team and Steering Committee, reviews the SoW to ensure it meets the CIE standards and selects the most qualified CIE reviewers according to the expertise requirements in the SoW. The CIE selection process also requires that CIE reviewers can conduct an impartial and unbiased peer review without the influence from government managers, the fishing industry, or any other interest group resulting in conflict of interest concerns. Each CIE reviewer is required by the CIE selection process to complete a Lack of Conflict of Interest Statement ensuring no advocacy or funding concerns exist that may adversely affect the perception of impartiality of the CIE peer review. The CIE reviewers conduct the peer review, often participating as a member in a panel review or as a desk review, in accordance with the ToR producing a CIE independent peer review report as a deliverable. The Office of Science and Technology serves as the COTR for the CIE contract with the responsibilities to review and approve the deliverables for compliance with the SoW and ToR. When the deliverables are approved by the COTR, the Office of Science and Technology has the responsibility for the distribution of the CIE reports to the Project Contact.

Requirements for CIE Reviewers:

Three CIE reviewers are requested to conduct an impartial and independent peer review in accordance with the Terms of Reference (ToR) herein. Each CIE reviewer's duties shall not exceed a maximum of 14 days conducting pre-review preparations with document review, participation on the SARC panel review meeting, editorial assistance to the SARC Chair, and completion of the CIE independent peer review report in accordance with the ToR and Schedule of Milestones and Deliverables. CIE reviewers shall have working knowledge and recent experience in the application of modern fishery stock assessment models. Expertise should include both the use of statistical catch-at-age and traditional VPA approaches. Experience with comparative studies of these approaches is especially valuable. Reviewers should also have experience in evaluating measures of model fit, identifiability, uncertainty, and forecasting. Some experience with groundfish (such as cod, haddock, flounder) population dynamics would be useful.

Specific Activities and Responsibilities

The CIE's deliverables shall be provided according to the schedule of milestones listed on page 5. The GARM Chair will use contributions from the CIE panelists as well as from other external panelists, to produce the GARM Panel Summary Report. In addition, each CIE panelist will write an individual independent report. These reports will provide peer-review information for a presentation to be made by NOAA Fisheries at meetings of the New England and Mid-Atlantic Fishery Management Councils in 2008. The GARM Panel Summary Report shall be an accurate and fair representation of the GARM panel viewpoint on the quality and soundness of the science, methods and data with regard to each Term of Reference (see Annex 1). The report shall also contain recommendations for improvement that might be implemented in a future GARM meeting.

Charge to GARM panel

The panel is to determine and write down its viewpoint on the quality and soundness of the science, methods and data with regard to each Term of Reference (see Annex 1). Criteria to consider include whether: (1) the data are adequate and were used properly; (2) the analyses and models were appropriate and correctly accomplished; and (3) the conclusions are correct/reasonable. Where possible, the chair shall identify or facilitate agreement among the panelists regarding each Term of Reference.

During the course of the review, the panel is allowed limited flexibility to deviate from the results and recommendations of earlier GARM-III meetings. This flexibility may include minor alterations in procedures previously established at the peer review of the Data Methods Meeting in October 2007. Large scale changes, such as changing a stock definition would not be possible in view of the difficulties of implementing these changes in time available before the final GARM meeting in August 2008.

Furthermore, if the panel rejects certain assessment models, the panel should explain why those particular models are not suitable, and the panel should recommend suitable alternatives. If such alternatives cannot be identified, then the panel should indicate that the existing (status quo) models are the best available at this time.

Roles and responsibilities

(1) Prior to the meeting

(GARM Chair and CIE panelists)

Review the reports produced by the Working Groups, and read background reports.

(2) During the Open meeting

(GARM Chair)

Act as chairperson, where duties include control of the meeting, coordination, control, and facilitation of the presentations and discussions, and ensuring that all Terms of Reference of the GARM are reviewed and completely addressed.

During the question and answer periods, provide appropriate feedback to the assessment scientists on the sufficiency of the analyses and when possible, suggest improved approaches. It is permissible to discuss the working papers, and to request additional information to clarify or revise existing analyses, if that information can be produced rather quickly.

(CIE panelists)

For each model approach, participate in panel discussions on the quality and soundness of the science, methods and data with regard to each Term of Reference (see Annex 1).

During the question and answer periods, provide appropriate feedback to the assessment scientists on the sufficiency of the analyses. It is permissible to request additional information if it is needed to clarify or revise existing analyses, if that information can be produced rather quickly.

(3) After the Open meeting

(GARM CIE panelists)

Each panelist shall prepare a CIE independent peer review report (see Annex 2). This report should comment on the quality and soundness of the science, methods, and data with regard to each Term of Reference.

If any modeling approaches are considered inappropriate, the CIE independent peer review report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing modeling approaches are the best available at this time.

During the meeting, additional questions that are not in the Terms of Reference but which are directly related to the assessments may be raised. Comments on these questions should be included in a separate section at the end of the CIE independent peer review report prepared by each panelist.

If a panelist feels that his/her comments are adequately expressed in the GARM Panel Summary Report, it will not be necessary to repeat the same comments in the Independent CIE Report. In this case, the CIE independent peer review report can be

used to provide greater detail on specific Terms of Reference or additional questions raised during the meeting.

(GARM Chair)
The GARM Chair shall prepare a document summarizing the background of the work to be conducted as part of the review process, and summarizing whether the process was adequate to successfully address the Terms of Reference. If appropriate, the chair will include suggestions on how to improve the process. This document will constitute the introduction to the GARM Panel Summary Report.

(GARM Chair, CIE and non-CIE panelists)
The GARM Chair will take the lead in preparing, editing, and completing the GARM Panel Summary Report, based on contributions from the external panelists (CIE and non-CIE). The panelists and the chair will discuss their views on each Term of Reference and whether their opinions can be summarized into a single conclusion for all—or only for some—of the Terms of Reference. For TORs where a consensus view can be reached, the GARM Panel Summary Report will contain a summary of such views. In cases where multiple and/or differing views exist on a given Term of Reference, the GARM Panel Summary Report will note that there was no agreement and will specify—in a summary manner—what the various opinions are and the reason(s) for the different opinions.

The Chair's objective during this Summary Report development process will be to identify or facilitate the finding of an agreement, rather than forcing the panel to reach an agreement if this is not possible.

The GARM Panel Summary Report (please see Annex 3 for information on contents) should comment on the quality and soundness of the science, methods, and data with regard to each Term of Reference.

If any modeling approaches are considered inappropriate, the GARM Panel Summary Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing modeling approaches are the best available at this time.

The contents of the draft GARM Panel Summary Report will be approved by the CIE panelists by the end of the Summary Report development process. The GARM chair will finalize all editorial and formatting changes prior to approval of the contents of the draft GARM Panel Summary Report by the CIE panelists. The GARM chair will then submit the approved GARM Panel Summary Report to the NEFSC contact (i.e., SAW Chair).

Schedule of Milestones and Deliverables

The milestones and schedule are summarized in the table below. No later than March 14, 2008, the CIE panelists should submit their CIE independent peer review reports to the CIE for review[1]. The CIE reports shall be sent to "University of Miami Independent System for Peer Review," and sent to Dr. David Sampson, via e-mail to David.Sampson@oregonstate.edu and to Mr. Manoj Shivlani via e-mail to mshivlani@ntvifederal.com

Milestone	**Date**
Open workshop at Northeast Fisheries Science Center (NEFSC) (begin writing reports, as soon as open Workshop ends)	Feb. 25 – 29, 2008
GARM Chair and CIE panelists work at the NEFSC drafting reports	Feb. 28 – 29
Draft of GARM Panel Summary Report, reviewed by all CIE panelists, due to the GARM Chair **	March 14
CIE panelists submit CIE independent peer review reports to CIE for approval	March 14
GARM Chair sends Final GARM Panel Summary Report, approved by CIE panelists, to NEFSC contact (i.e., SAW Chairman)	March 21
CIE provides reviewed CIE independent peer review reports to NMFS COTR for approval	March 28
COTR notifies CIE of approval of CIE independent peer review reports	April 4 *
COTR provides final CIE independent peer review reports to NEFSC contact	April 4

* Assuming no revisions are required of the reports.
** The GARM Panel Summary Report will not be submitted, reviewed, or approved by the CIE.

The SAW Chairman will assist the GARM chair prior to, during, and after the meeting in ensuring that documents are distributed in a timely fashion. NEFSC staff and the SAW Chairman will make the final GARM Panel Summary Report and CIE independent peer review reports available to the public. Staff and the SAW Chairman will also be responsible for production and publication of the collective Working Group papers.

Acceptance of Deliverables:

Upon review and acceptance of the CIE reports by the CIE Coordination and Steering Committees, CIE shall send via e-mail the CIE reports to the COTRs (William Michaels William.Michaels@noaa.gov and Stephen K. Brown Stephen.K.Brown@noaa.gov) at the NMFS Office of Science and Technology by the date in the Schedule of Milestones and Deliverables. The COTRs will review the CIE reports to ensure compliance with the SoW and ToR herein, and have the responsibility of approval and acceptance of the deliverables. Upon notification of acceptance, CIE shall send via e-mail the final CIE report in *.PDF format to the COTRs. The COTRs at the Office of Science and Technology have the responsibility for the distribution of the final CIE reports to the Project Contacts.

[1] All reports will undergo an internal CIE review before they are considered final.

Key Personnel:

Contracting Officer's Technical Representative (COTR):

William Michaels
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
William.Michaels@noaa.gov Phone: 301-713-2363 ext 136

Stephen K. Brown
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
Stephen.K.Brown@noaa.gov Phone: 301-713-2363 ext 133

Contractor Contacts:

Manoj Shivlani, CIE Primary Coordinator
10600 SW 131st Court, Miami, FL 33186
mshivlani@ntvifederal.com Phone: 305-383-4229

Roger Peretti, NTVI Regional Director
Northern Taiga Ventures, Inc., 814 W. Diamond Ave., Ste. 250, Gaithersburg, MD 20878
rperetti@ntvifed.com Phone: 301-212-4187

Project Contact:

James Weinberg, NEFSC Contact person and SAW Chairman
NMFS Northeast Fisheries Science Center, 166 Water Street, Woods Hole, MA 02543
James.Weinberg@noaa.gov Phone: 508-495-2352

Request for Changes:

Requests for changes shall be submitted to the Contracting Officer at least 15 working days prior to making any permanent substitutions. The Contracting Officer will notify the Contractor within 10 working days after receipt of all required information of the decision on substitutions. The contract will be modified to reflect any approved changes. The Terms of Reference (ToR) and list of pre-review documents herein may be updated without contract modification as long as the role and ability of the CIE reviewers to complete the SoW deliverable in accordance with the ToR are not adversely impacted.

ANNEX 1:

Draft Terms of Reference for the GARM-III "Models" Meeting

(Last Revised: Oct. 31, 2007; A final draft will be distributed to the Panel prior to the meeting.)

1. For each stock, consider the applicability of one or more of the following modeling approaches to assess stock status:
    - Index methods
    - Production Models
    - Age- or Length-based Models

2. For certain stocks that are aged, compare and contrast the utility of statistical catch-at-age vs. VPA based models with respect to the following criteria:
    - Retrospective patterns
    - Flexibility to account for alternative parameterizations
    - Ability to incorporate external sources of information, especially tagging and environmental data
    -Ability to estimate parameters incorporating prior, external information.

3. Address the implications of zeros in the evaluation of fishery independent indices.

4. Examine potential factors responsible for retrospective patterns.

5. For each stock, define the assessment model that will be used to determine stock status and productivity characteristics until the next "benchmark" assessment is conducted. Where possible, apply the models to data (probably through 2006), to obtain current and historical estimates of F and B and estimates of uncertainty.

6. Evaluate the sufficiency of the assessment models to estimate measures of stock status consistent with Biological Reference Points.

ANNEX 2: Contents of GARM-III CIE independent peer review report

1. The Independent CIE Report should comment on the quality and soundness of the science, methods and data with regard to each Term of Reference. CIE panelists should consider whether the work provides a scientifically credible basis for developing fishery management advice. Scientific criteria to consider include: whether the data were adequate and used properly, the analyses and models were carried out correctly, and the conclusions are correct/reasonable.

If a panelist feels that his/her comments are adequately expressed in the GARM Panel Summary Report, it will not be necessary to repeat the same comments in the Independent CIE Report. In that case, the Independent CIE Report can be used to provide greater detail on specific Terms of Reference or additional questions raised during the meeting.

2. If any modeling approaches are considered inappropriate, the Independent CIE Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing modeling approaches are the best available at this time.

3. Any independent analyses conducted by the CIE panelists as part of their responsibilities under this agreement should be incorporated into their Independent CIE Reports. It would also be helpful if the details of those analyses (e.g., computer programs, spreadsheets etc.) were made available to the respective assessment scientists.

4. Additional questions that were not in the Terms of Reference but that are directly related to the assessments should be addressed. This section should only be included if additional questions were raised during the GARM meeting.

ANNEX 3: Contents of GARM-III Panel Summary Report

1. The first section the report shall consist of an introduction prepared by the GARM chair that will include the background, a review of activities and comments on the appropriateness of the process in reaching the goals of the GARM. The next section will contain comments on the quality and soundness of the science, methods and data with regard to each Term of Reference. The GARM Panel should consider whether the work provides a scientifically credible basis for developing fishery management advice. Scientific criteria to consider include: whether the data were adequate and used properly, the analyses and models were carried out correctly, and the conclusions are correct/reasonable.

If the CIE panelists, the non-CIE panelists and GARM chair do not reach an agreement on a Term of Reference, the report should explain why. It is permissible to express majority as well as minority opinions.

2. If any modeling approaches are considered inappropriate, the GARM Panel Summary Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing modeling approaches are the best available at this time.

3. The report shall also include the bibliography of all materials provided during the meeting and any papers cited in the GARM Panel Summary Report, along with a copy of the CIE Statement of Work.

The report shall also include as a separate appendix the Terms of Reference used for the GARM Models Meeting, including any changes to the Terms of Reference or specific topics/issues directly related to the assessments and requiring Panel advice.